Computational analysis and generation of the nominal and verbal complexes

Dr. S.Rjendran
Department of Linguistics
Tamil University
Thanjavur 613 005

Computational analysis of nominal verbal complexes involves two processes:

1. The analysis of nominal and verbal complex into minimal meaningful units and assigning semantic features to each unit.

1. The generation of the nominal and verbal complex out of the same meaningful units.

Let us make an assumption that the building of nominal and verbal complexes is the concatenation of morphemes and bound words. Clearly the simplest way to build such a model would be to assume that the allowability of a particular morpheme or a bound word in a given context depends only upon the morpheme that precedes it. Take the word *ceytuviTTeen*, whose morphological structure is *cey+tu + viT+T+een*. The combination of *cey* which is a verb and *tu* which an adverbial marker forms the first chunk of the verbal complex and the combination of *viT* which is a secondary verb, *T* which is a past tense marker and *een* which is a person-number-marker forms the second chunk of the word. Both of the chunks (V1 and V2) make a verbal complex.

Finites state Automata

One can easily model such a system with a finite-state machine (finite-state automaton or finite-state transition network) of the kind familiar from formal language theory. More formally, a finite automation is defined by the following five parameters:

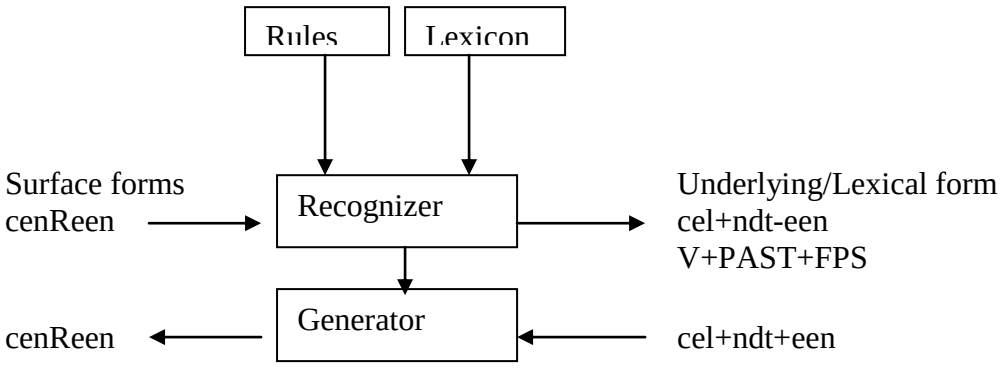
Q: a finite set of N states q0, q1...,qN
Σ: a finite input alphabet of symbols
q0: the start state
F: the set of final states, $F \subseteq Q$
δ (q,i): the transition function or transition matrix between states. Given a state $q \in Q$ and an input symbol $i \in \Sigma$, δ (q,i) returns a new state $q' \in Q$. δ is thus a relation from $Q \times \Sigma$ to Q;

A set of rewrite rules, or grammar, can be thought of not only as an intentional definition of the units of a language, but also as the description of an abstract machine which is able to accept (or generate) just those strings which belong to the language. These machines are called (abstract) automata, and the complexity of their potential behaviour is directly determined by the internal complexity of the rules of the grammar to which they correspond. All such machines may be characterized in terms

of a set of states, an input device which can access one input symbol at a time, and control unit which can examine and read input and cause the machine to shift from one sate to another. In addition, such machines may have memory, which can be accessed by the control unit for storing and testing symbols. The computational power of the machine is determined essentially by the complexity of its memory and of the operations that can be performed on the memory. The simplest of these automata are called finite state machines or finite state automata (FSA), and the languages they can accept are regular languages.

PCKimmmo

The two level morphology of Kosekenniemi has been a widely implemented one, particularly in European languages. PCKimmo can do two functions as a morphological processor, recognition. In the case of the generator component of PCKimmo, it takes a lexical form of the foot as input. To this imput rules recorded in the rules file are applied and the corresponding surface form is returned. It does not use the lexicon. The recognizer accepts as input a surface form, applies the rules, consults the lexicon and returns the corresponding lexical form and gloss string. What makes it easier for the user and important from the point of economy of model is the fact that the same set of rules can be used for both the functions of recognition and gneration. The rules can be used birectionally (1990:8). If certain phonological, or of for that matter a morphological or orthographic, rule is written to refer to a transformation form the underlying structure to surface structure, they need not be rewritten for the reverse process. In reverse, it processes form the surface form to the underlying form. This model can be represented through the following diagram (Simons 1989).



(FPS=First person singular)

The two levels used in this model are linguistic descriptions that can handle any tow levels of abstractions, be it allophonic alternations, morphophonemic alternations or orthographic variations.

API Morphological Analyzer

The API Morphological analyzer developed by AUKBC Research Centre in collaboratin with the author of the paper depends on automata table present in its data file, statetable.dat. The API morphological Analyzer is actually designed to strip a word of its inflections.

The machine while processing is set to be at the state being represented by 0:0. Then the input is looked into root dictionary for a match. If a match is found it returns it along with its gloss and by printing, 'FINISH' in the command line. If it doesn't find the input in the root file, then it starts matching characters form the right periphery with the entries listed in the file 'statetable.dat.' If it doesn't find a match there then it goes and checks in a file called 'SANDHI.dat'. If it finds a morpheme, then it strips the word of the morpheme, takes the remaining part of the word to be checked again in a file called 'ROOT.dat'. The sandhi replacement tables are in the file 'sandhidict.dat'. For all these operations the analyzer depends on automata tables present in the data file 'statetable.dat'. This table contains various rules defining the paths needed to identify the affixes in a Tamil word. The state table is based on the verb flow chart. The format in which the rules are represented in this file is as fallows:

Stateno:morpheme,next state [-nextstate...]; [morpheme, nextstate];

An example for a morpheme in the suffix file will be as follows:

Ex.3.0:avan,1.1;avaL,1.1; avar,1.1;atu,1.1;avai,1.1; ana,1.1; kai,2.1.;tal,2.1; #ttal,2.1-End;Ral,2.1.-End

The fist entry in the line (stateno) is an arbitrary number (or word) designating the state that the automata must be in when it scans the line. This 'stateno' is separated form other entries by a colon. Any number of entries may follow the 'stateno'. A semicolon terminates each entry. The semicolon is only optional for the last entry. Each entry of two fields, the morpheme and the number of the next state that the automata will transit into if the morpheme in the entry is present at the end of word. These two fields are separated by a coma.

The following is an example for a rule in the file which strips the word the change because of the sandhi operations.

Ex. maram+ai==> marattai

The correspondence shows all that *m* will be replaced by *tt* when *ai* is added. During analysis *tt* will be replaced by the original *m* deletion. The entry in the rule file will be in the following format.

Ex.
4a +/-
<replacementCharacter>, {{Suffix{+postfix:[postfix...]};
{Suffix{+postfix:[postfix...]};

The morpheme entry will be as follows:

Ex.4b. –m, {tt{+ai:+aal:+uTan:+in:+ttin:+il:}; ng{+kaL:}}

The '-' sign before *m* indicates that the character will be added. The suffixes are enclosed in curly brackets. In the case of multiple suffixes, as is the case above, semicolons separate each suffix.

Different modules are used to do different functions. The modules have been written in the compiled language Java. The main functions load is done by TamilMorph.java. A module called RootDict. java does the search for the match in the root file. Similarly Suffix.java, Sandhi.java etc. take care of various functions using the finite state automata.

Noun Classes

class-1 *kaTaa*-class, class-2 *eli*-class, class-3 *ii*-class, class-4 *vaNTu* class, class-5 *kaaTu*-class, class-6 *aaRu*-class, class-7 *kaN* class, class-8 *pon*-class, class-9 *ndey*-class, class-10 ndaay-class, class-11 *pul*-class, class-12 *muL*-class, class-13 *maram*-class, class 14 *pala*-class, class-15 *ellaam*-class, class-16 *cuvar*-class, class-17 *kaal*-class, 18-class *ndaan*-class, class-19 *ndaam*-class, class-20 *ndaangkaL*-class, class-21 *ndii*-class, class-22 *ndiir* class, class-23 *ndiingkaL*-class, class 24 *taan*-class, class-25 *taangaL*-class, class-26 *atu*-class, class-27 *avai*-class.

Verb Classes

class-1 *cey*-class, class-2 *azu*-class, class-3 *caa*-class, class-4 *viTu*-class, class-5 *peRu*-class, class-6 *pukaz*-class, class-7 *vizu*-class, class-8 *aRi*-class, class-9 *ndoo*-class, class-10 *vaa*-class, class-11 *akal*-class, class-12 *kol*-class, class-13 *aaL*-class, class-14 *koL*-class, class-15 *ooTu*-class, class-19 *keeL*-class, class-20 *puuN*-class, class-21 *uN*-class, class-22 *kaaN*-class, class-23 *tin*-class, class-24 *ndil*-class, class-25 *paTi*-class, class-26 *ndaTa*-class

Tabulation of the flow of morphemes

Tables for Nominal complex

The tables given below show the flow of suffixes form one level to another level and the set of suffixes occurring in each level.

Table 1 showing the first level base and second level of suffixation

Noun class	**Base alternatns**	**Plural suffix**	**Accusative case suffix**	**Instrumental case suffix**	**Dative case suffix**	**Genitive case suffix-1**	**Genitive case suffix-2**	**Genitive case suffix-3**	**Locative case suffix**	**Sociative case suffix**	**Sociative case suffix**
class-1 *kaTaa*-class	kaTaa ~ kaTaav	kkaL	ai/inai	aal/inaal	ukku/inukku	in	atu/inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class-*2 eli*-class	eli ~ eliy	kaL	ai/inai	aal/inaal	ukku/inukku	in	atu/inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class 3 *ii*-class	ii~iiy	kkaL	ai/inai	aal/inaal	ukku/inukku	in	atu/inatu	uTaiya/inuTaiya uTaiya/inuTaiya	il/inil	ooTu	uTan
class 4 *vaNTu*-class	vaNTu ~ vaNT	kaL	ai/inai	aal/inaal	ukku/inukku	in	atu/inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class 5 *kaaTu*-class	kaaTu ~ kaaTT	kaL	ai/inai	aal/inaal	ukku/inukku	in	atu/inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class 6 *aaRu*-class	aaRu ~aaRR	aaRukaL	ai/inai	aal/inaal	ukku/inukku	in	atu/inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class 7 *kaN*-class	kaN ~ kaNN	kaL	ai/inai	aal/inaal	ukku/inukku	in	atu/inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class 8 *pon*-class	pon ~ ponn	kaL	ai/inai	aal/inaal	ukku/inukku	in	atu/inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class 9 *ndey*-class	ndey ~ ndeyy	kaL	ai/inai	aal/inaal	kku/inukkku	in	atu/inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class 10 *ndaay*-class	ndaay	kaL	ai/inai	aal/inaal	kku/inukkku	in	atu/inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class 11 *pul*-class	pul ~ pull~puR	kaL	ai/inai	aal/inaal	ukku/inukku	in	atu/inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class 12	muL ~ muLL	kaL	ai/inai	aal/inaal	ukku/inuk	in	atu/inatu	uTaiya/inuT	il/inil	ooTu	uTan

***muL*-class**	~muT				ku			aiya			
class 13 *maram*-class	maram ~ marang~ maratt	kaL	ai/inai	aal/inaal	ukku /inukku	in	atu/ inatu	uTaiya/inuTaiya il/inil	il/inil	ooTu	uTan
class 14 *pala*-class	pala ~ palavaRR	-	ai/inai	aal/inaal	ukku /inukku	in	atu/ inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class 15 *ellaam*-class	ellaam ~ ellaavaRR	-	ai/inai	aal/inaal	ukku /inukku	in	atu/ inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class 16 *cuvar*-class	cuvar ~cuvaRR	kaL	ai/inai	aal/inaal	ukku /inukku	in	atu/ inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class 17 *kaal*-class	kaal	kaL	ai/inai	aal/inaal	ukku /inukku	in	atu/ inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class 18 *ndaan*-class	ndaan ~en~ enn	-	ai/inai	aal/inaal	ukku /inukku	in	atu/ inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class 19 *ndaam*-class	ndaam ~ndam ~ndamm	-	ai/inai	aal/inaal	ukku /inukku	in	atu/ inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class 20 *ndaangkaL*-class	ndaangkaL~ engkaL	-	ai/inai	aal/inaal	ukku /inukku	in	atu/ inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class 21 *ndii*-class	ndii~ un~ unn	-	ai/inai	aal/inaal	ukku /inukku	in	atu/ inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class 22 *ndiir*-class	ndiir~ um~ umm	-	ai/inai	aal/inaal	ukku /inukku	in	atu/ inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class 23 *ndiingaL*-class	ndiingkaL~ ungkaL	-	ai/inai	aal/inaal	ukku /inukku	in	atu/ inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class 24 *taan*-class	taan~ tan~ tann	-	ai/inai	aal/inaal	ukku /inukku	in	atu/ inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class 25 *taangkaL*-class	taangkaL~tangkaL	-	ai/inai	aal/inaal	ukku /inukku	in	atu/ inatu	uTaiya/inuTaiya	il/inil	ooTu	uTan
class	avai~	-	ai/inai	aal/ina	ukku	in	atu/	uTaiy	il/inil	ooTu	uTan

26 *avai*-class	avaRR			al	/inuk ku		inatu	a/inuT aiya			

Table 2 showing the second/third level of

Set of postpositions occurring immediately after the nominative forms	**Set of postpositions occurring immediately after the oblique forms or after the *in* suffixed oblique forms**	**Set of postpositions occurring after the accusative case suffix *ai***	**Set of postpositions occurring after the dative case suffix *ukku***	**The postposition occurring after the locative case suffix *il* and other star (*) marked postpositions**
uTan	iTam*	paTTi	pakkam*	irundtu
varai	pakkam*	kuRittu	pakkattil*	
varaikkum	pakkattil*	oTTi	aNTai*	
illaamal	aNTai	paarttu	aNTaiyil*	
muulam	aNTaiyil*	ndookki	arukee*	
aaka	arukee*	cuRRi	arukil*	
aana	arukil*	taaNTi	aTuttaaRpoola*	
	aTuttaaRpoola	tavirttu	kiTTee*	
	kiTTee*	tavira	meel*	
	meel*	oziya	meelee*	
	meelee*	koNTu	kiiz*	
	kiiz*	vaittu	kiizee*	
	kiizee*	viTTu	uL*	
	uL*	poola	uLLee*	
	uLLee	maatiri	veLiyee*	
	veLiyee*	viTa	veLiyil*	
	veLiyil*		mun*	
	mun*		munnaal*	
	munnaal*		pinnaal	
	pinnaal		pinpu*	
	pinpu*		pindti	
	pindti		kuRukkee*	
	etiree*		etiree*	
	etiril*		etiril*	
	etirttaaRpool*		etirttaaRpool*	
	iTaiyil*		iTaiyil*	
	ndaTuvil*		ndaTuvil*	
	kuRukkee*		piRaku*	
			appuRam*	
			appaal*	
			vaTakku	
			vaTakkee*	
			vaTakkil*	
			kizakku	
			kizakkee*	
			kizakkil*	
			teRku	
			teRkee*	
			teRkuil*	
			meeRku	
			meeRkee*	
			aaka	
			veeNTi	

			patil	
			patilaaka	

('*' indicates the occurrence of *irundtu* after this postpositions.)

Tables for Verbal complex

Table 1: The vebal bases occurring at the first level and the suffixes occurring at the second level

Verb class	**Stem alternants**	**past tense suffix (occurring immediately after verb stem)**	**present tense suffix (occurring immediately after verb base)**	**future tense suffix (occurring immediately after verb base)**	**negative suffix (occurring immediately after verb base)**	**Infinitive suffix (occurring immediately after verb base)**	**Adverbial suffix (occurring immediately after verb base)**	**Nominizer 1 (occurring immediately after the verb)**	**Nominizer 2 (occurring immediately after the verb)**	**Nominer (occurring immediately after t verb)**
class 1 *cey* class	cey ~ ceyy	t	kiR ~ kinR	v ~ um	aa~aat	a	tu ~t	tal	al	kai
class 2 *azu* class	azu ~ az	t	kiR ~ kinR	v ~ um	aa~aat	a	tu~t	tal	al	kai
class 3 *caa* class	caa ~ ce	t	kiR ~ kinR	v ~ kum	kaa~kaat	ka	ttu~tt	tal	kal	kai
class 4 *viTu* class	viTu ~ viT	T	kiR ~ kinR	v ~ um	aa~aat	a	Tu~T	tal	al	kai
class 5 *peRu* -class	peRu~ peR	R	kiR ~ kinR	v ~ um	aa~aat	a	Ru~R	tal	al	kai
class 6 *pukaz* -class	pukaz	ndt	kiR ~ kinR	v ~ um	aa~aat	a	ndtu~ ndt	tal	al	kai
class 7 *vizu*- class,	vizu~ viz	ndt	kiR ~ kinR	v ~ um	aa~aat	a	ndtu~ ndt	tal	al	kai
class 8 *aRi*- class	aRi~ aRiy	ndt	kiR ~ kinR	v~um	aa~aat	a	ndtu~ ndt	tal	al	kai
class 9 *ndoo*- class	ndoo~ ndo~	ndt	kiR ~ kinR	v~kum	kaa~kaat	ka	ndtu~ ndt	tal	kal	kai
class 10 *vaa*- class,	vaa~ va~ varu~ var	ndt	kiR ~ kinR	v~um	aa~aat	a	ndtu~ ndt	tal	al	kai
class	akal~	nR	kiR ~	v~um	aa~aat	a	nRu~	tal	al	kai

11 *akal*-class	aka		kinR				nR			
class 12 *kol*-class	kol ~ko ~ koll	nR	kiR ~ kinR	v~um	aa~aat	a	nRu~ nR	tal	al	kai
class 13 *aaL*-class	aaL~aa	NT	kiR ~ kinR	v~um	aa~aat	a	NTu ~NT	tal	al	kai
class 14 *koL*-class	koL~ko ~ koLL	NT	kiR ~ kinR	v~um	aa~aat	a	NTu~ NT	tal	al	kai
class 15 *ooTu*-class	ooTu~ ooT	in	kiR ~ kinR	v~um	aa~aat	a	i~iy	tal	al	kai
class 16 *col*-class	col~co~ coll	nn	kiR ~ kinR	v~um	aa~aat	a	i~iy	tal	al	kai
class 17 *poo*-class	poo	n	kiR ~ kinR	v~kum	kaa~kaat	ka	y	tal	kal	kai
class 18 *kal*-class	kal~ kaR	R	kiR ~ kinR	p~kum	kaa~kaat	ka	Ru~R	Ral	kal	kai
class 19 *keeL*-class	keeL~ keeT	T	kiR ~ kinR	p~kum	kaa~kaat	ka	Tu~T	tal	kal	kai
class 20 *puuN*-class	puuN	T	kiR ~ kinR	p~um	aa~aat	a	Tu~T	tal	al	kai
class 21 *uN*-calss	uN~uN N	T	kiR ~ kinR	p~um	aa~aat	a	Tu~T	tal	al	kai
class 22 *kaaN*-class	kaaN~ kaN	T	kiR ~ kinR	p~um	aa~aat	a	NTu	utal	al	kai
class 23 *tin*-class	tin ~tinn	R	kiR ~ kinR	p~um	aa~aat	a	nRu~ nR	utal	al	kai
class 24 *ndil*-class	ndil~ ndi~ ndiR ~ndill	nR	kiR ~ kinR	p~kum	kaa~kaat	ka	nRu	Ral	kal	kai
class 25 *paTi*-class	paTi	tt	kkiR ~ kkinR	pp~ kkum	kkaa~ kkaat	kka	ttu~ tt	ttal	kkal	kkai

class 26 *ndaTa*-class	ndaTa	ndt	kkiR ~ kkinR		kkaa~ kkaaat	kka	ndtu~ ndt	ttal	kkal	kkai

Table 2: The suffixes occurring at the third and fourth level

Verb class	**PNG suffix (occurring immediately after tense suffix)**	**Neuter PNG suffix (occurring immediately after the negative suffix *aa*)**	**Adverbial suffix (occurring immediately after negative *aa*)**	**Adjectival suffix (occurring immediately after tense suffix or negative suffix *aat*)**	**Conditional suffix (occurring immediately after past tense suffix)**	**Nominalizer 4 (occurring immediately after the negative *aa*)**	**Gerundival nominalizers occurring after adjectival participle)**	**Pronominalizers (occurring immediately after the adjectival participle)**
class 1 *cey* class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
class 2 *azu* class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
class 3 *caa* class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	
class 4 *viTu* class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
class 5 *peRu*-class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a		mai	tu	van/vaL/var/ varkaL/tu/vai
class 6 *pukaz*-class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
class 7 *vizu*-class,	een/oom/aay /iir/iirkaL/aan/ aaL/aar/	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai

	atu/ana							
class 8 *aRi-* class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
class 9 *ndoo-* class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
class 10 *vaa-* class,	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
class 11 *akal-* class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
class 12 *kol-* class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
class 13 *aaL-* class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
class 14 *koL-* class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
class 15 *ooTu-* class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
class 16 *col-* class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
class 17 *poo-* class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai

	atu/ana							
class 18 *kal*- class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
class 19 *keeL*- class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
class 20 *puuN*- class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
class 21 *uN*- calss	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
class 22 *kaaN*- class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
class 23 *tin*- class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
class 24 *ndil*- class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
class 25 *paTi*- class	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana een/oom/aay /iir/iirkaL/aan/ aaL/aar/	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai

	atu/ana							
class 26 *ndaTa*-class	een/oom/aay /iir/iirkaL/ aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai

Table 3: The third level auxiliary verbs occurring after infinitive and the suffixes of second occurring at the fourth level

Auxilary verbs occurring after infinitive form	**Past tense suffix (occurring immediately after verb stem)**	**Present tense suffix (occurring immediately after verb base)**	**Future tense suffix (occurring immediately after verb base)**	**Negative suffix (occurring immediately after verb base)**	**Infinitive suffix (occurring immediately after verb base)**	**Adverbial suffix (occurring immediately after verb base)**	**Nominizer 1 (occurring immediately after the verb)**	**Nominizer 2 (occurring immediately after the verb)**	**Nominizer 1 (occurring immediately after the verb)**
TTum *	-	-	-	-	-	-	-	-	-
muTi~ muTiy	ndt	kiR ~ kinR	um	aa~aat	a	ndtu~ ndt	-	-	-
iyal	-	-	um	aa~aat	a	-	-	-	-
kuuTum*	-	-	-	-	-	-	-	-	-
kuuT	-	-	um	aa~aat	-	i	-	-	-
paar	tt	kkiR ~ kkinR	pp~ kkum	kkaa~ kkaat	kka	ttu~ tt	ttal	kkal	kkai
iru	ndt	kkiR ~ kkinR	pp ~kkum	aa~aat	kka	ndtu~ndt	ttal	kkal	kkai
poo	n	kiR ~ kinR	v~kum	kaa~kaat	ka	y	tal	kal	kai
varu~var~ va~vaa (1)	ntd	kiR ~ kinR	v~um	aa~aat	a	nRu~ nR	tal	al	kai
varu~var~ va~vaa (2)	-	kiR~kinR	um	aa~aat	a	ndtu~ ndt	-	-	-
veeNT	-	-	um	-	-	i ~iy	-	-	-
veeNTivaru ~ va	ndt	kiR~kinR	um	aa~aat	a	ndtu ~ndu	tal	al	kai
maaTT	-	-	-	-	-	-	-	-	-
illai *	-	-	-	-	-	-	-	-	-
cey~ceyy	t	kiR ~ kinR	v ~um	aa~aat	a	tu~t	tal	al	kai
vai	tt	kkiR ~ kinR	pp ~ kkum	kkaa ~ kkaat	a	ttu ~ t	ttal	kkal	kkai
paTu ~ paT	T	kiR ~ kinR	v ~ um	aa ~ aat	a	Tu ~ T	tal	al	kai

Table 4: The third level auxiliary verbs occurring after infinitive and the suffixes of third level occurring at the fifth level and the suffixes of fourth level occurring at the sixth level

Auxilary verbs	**PNG suffix**	**Neuter PNG**	**Adverbial suffix**	**Adjectival suffix**	**Conditional suffix**	**Nominalizer 4**	**Gerundival**	**Pronominalizers**

occurring after infinitive form	**(occurring immediately after tense suffix)**	**suffix (occurring immediately after the negative suffix *aa*)**	**(occurring immediately after negative *aa*)**	**(occurring immediately after tense suffix or negative suffix *aat*)**	**(occurring immediately after past tense suffix)**	**(occurring immediately after the negative *aa*)**	**nominalizers occurring after adjectival participle)**	**(occurring immediately after the adjectival participle)**
Ttum *	-	-	-	-	-	-	-	-
muTi~ muTiy	-	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
iyal	-	tu	mal~tu	-	-	mai	tu	van/vaL/var/ varkaL/tu/va
kuuT		tu	-	a	-	-		van/vaL/var/ varkaL/tu/vai
paar	een/oom/aay /iir/iirkaL/ aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
iru	een/oom/aay /iir/iirkaL/ aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal		tu	van/vaL/var/ varkaL/tu/vai
poo	een/oom/aay /iir/iirkaL/ aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
varu~var~va~vaa (1)	een/oom/aay /iir/iirkaL/ aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu/vai	van/vaL/var/ varkaL/tu/vai
varu~var~va~vaa (2)	een/oom/aay /iir/iirkaL/ aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
veeNT-	-	-	-	a	-	-	tu	van/vaL/var/ varkaL/tu/vai
maaTT	een/oom/aay /iir/iirkaL/ aan/ aaL/aar/	tu	-	-	-	-	tu	van/vaL/var/ varkaL/tu/vai

	atu/ana							
illai *	-	-	-	-	-	-	-	-
cey~ceyy	een/oom/aay /iir/iirkaL/ aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var / varkaL/tu/vai
vai	een/oom/aay /iir/iirkaL/ aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	

Table 5: The third level auxiliary verbs occurring after adverbial participle and the suffixes of second level occurring at the fourth level

The auxiliary verbs occurring after adverbial participle	**past tense suffix (occurring immediately after verb stem)**	**present tense suffix (occurring immediately after verb base)**	**future tense suffix (occurring immediately after verb base)**	**negative suffix (occurring immediately after verb base)**	**Infinitive suffix (occurring immediately after verb base)**	**Adverbial suffix (occurring immediately after verb base)**	**Nominizer 1 (occurring immediately after the verb)**	**Nominizer 2 (occurring immediately after the verb)**	**Nominizer 3 (occurring immediately after the verb)**
aayiRRu *									
iru (class	ndt	kkiR~kkinR	pp~kkum	kkaa	a	ndtu ~ndt	ttal	al	kkai
viTu	T	kiR~kinR	v~um	aa	a	Tu~T	tal	al	kai
poo	n	kiR~kinR	v~kuu	kaa	a	y	tal	kal	kai
tiir	tt	kiR~kiR	pp~kkum	aa	ka	ttu~tt	ttal	kkal	kkai
muTi	tt	kkiR	pp~kkum	kkaa	kka	ttu~tt	ttal	kkal	kkai
pooTu	T	kiR	v~um	aa	a	Tu~T	tal	al	kai
koNTiru	ndt	kkiR	pp~kkum	kkaa	kka		tta	kkal	kkai
koL	NT	kiR	v~	aa	a		tal	al	kai
paar	tt	kkiR~kkinR	pp~kum	kkaa	kka	ttu~tt	ttal	kkal	kkai
koTu	tt	kkiR~kinR	pp~kum	kkaa	kka	ttu~tt	ttal	kkal	kkai
tolai	tt	kkiR~kinR	pp~kum	kkaa	kka	ttu~tt	ttal	kkal	kkai
kizi	tt	kkiR~kinR	pp~kum	kkaa	kka	ttu~tt	ttal	kkal	kkai
taLLu	in	kiR~kin	v~um	aa	a	i	tal	al	kai

Table 6: Table 5: The third level auxiliary verbs occurring after adverbial participle and the suffixes of third level occurring at the fifth level and the suffixes of fourth level occurring at the sixth level

The auxiliary verbs occuring after	**PNG suffix (occurring immediat**	**Neuter PNG suffix (occurring**	**Adverbial suffix (occurring immediat**	**Adjectival suffix (occurring immediat**	**Conditional suffix (occurring immediat**	**Nominalizer 4 (occurring immediat**	**Gerundival nominalizers occurring**	**Pronominalizers (occurring immediately after the**

adverbial participle	**ely after tense suffix)**	**immediat ely after the negative suffix *aa*)**	**ely after negative *aa*)**	**ely after tense suffix or negative suffix *aat*)**	**ely after past tense suffix)**	**ely after the negative *aa*)**	**after adjectival participle)**	**adjectival participle)**
aayiRRu *	-	-	-	-	-	-	-	-
iru	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal		tu	van/vaL/var/ varkaL/tu/vai
viTu	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a		mai	tu	
poo	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
tiir	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a		mai	tu/vai	van/vaL/var/ varkaL/tu/vai
muTi	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
pooTu	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
koNTiru	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
koL	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
paar	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai

koTu	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
tolai	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
kizi	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai
taLLu	een/oom/aay /iir/iirkaL/aan/ aaL/aar/ atu/ana	tu	mal~tu	a	aal	mai	tu	van/vaL/var/ varkaL/tu/vai

Table 7: The auxiliary verbs occuring after *al*-nominalized form

Verbs	**past tense suffix (occurring immediately after verb stem)**	**present tense suffix (occurring immediately after verb base)**	**future tense suffix (occurring immediately after verb base)**	**negative suffix (occurring immediately after verb base)**	**Infinitive suffix (occurring immediately after verb base)**	**Adverbial suffix (occurring immediately after verb base)**	**Nominizer 1 (occurring immediately after the verb)**	**Nominizer 3 (occurring immediately after the verb)**	**Nominizer 3 (occurring immediately after the verb)**
aam#									
aakum#									
aakaatu#									

Table 8: The auxiliary verb occuring after *atu*-nominalized firm

Verbs				
uNTu#				
illai#				
kiTaiyaatu#				

(The marker # in the above tables denotes that the verbal forms are the finite forms and cannot be inflected further.)

Conclusion

The preparation of morphological analyzer is the preliminary step to be made before venturing into any other type of natural language processing. Tamil being a complex inflectional language is a challenge for computational linguists who try to prepare a morphological analyzer for it. There are attempts to prepare morphological generators and analyzers for Tamil. Preparing a morphological generator or analyzer for the purpose of demonstration is quite different from preparing an exhaustive and efficient working system. It is not too difficult to prepare a system with the coverage of 60 to 65%. However increasing the efficiency beyond 65% and to reach a realistic level of 95% to 97% is a hard task. Most of the existing descriptions focus on a narrow and vertical detail rather than on breadth and exhaustiveness. Generally the coverage obtained by incorporating the descriptions of the existing work is never beyond 50%. The evaluation method points out that the efficiency of the morphological analyzer prepared for Tamil by the present team is very remarkable.

The preparation of a morphological analyzer, in which the processing of verbal complex forms a part, has may natural language applications such as parsing, text generation, machine translation, preparing dictionary tools and lemmatization. It also helps in speech applications such as text-to-speech synthesizing and speech recognition and in word processing applications like spell checking and text input and in retrieval of documents. Finding the categorical details of the verbal forms helps in assigning parts-of-speech taggs to the verbal complex.